

Appln. No.: 09/820,401  
Amendment Dated March 10, 2004  
Reply to Office Action of December 11, 2003

MATP-601US

**Amendments to the Claims:** This listing of claims will replace all prior versions, and listings, of claims in the application

Listing of Claims:

---

1. (Currently Amended) A method of displaying text information corresponding to a speech portion of audio signals of a television program ~~te~~ as a closed caption on an video display device, the method comprising the steps of:

decoding the audio signals of the television program;

filtering the audio signals by using a spectral subtraction method to extract the speech portion;

parsing the speech portion into discrete speech components in accordance with a speech model and grouping the parsed speech components;

identifying words in a database corresponding to the grouped speech components; and

converting the identified words into text data for display on the display device as the closed caption.

2. (Original) A method according to claim 1, wherein the step of filtering the audio signals is performed concurrently with the step of decoding of later-occurring audio signals of the television program and step of parsing of earlier occurring speech signals of the television program.

3. (Original) A method according to claim 1, wherein the step of parsing the speech portion into discrete speech components includes the step of employing a speaker independent model to provide individual words as the parsed speech components.

4. (Original) A method according to claim 1 further including the step of formatting the text data into lines of text data for display in a closed caption area of the display device.

Appln. No.: 09/820,401  
Amendment Dated March 10, 2004  
Reply to Office Action of December 11, 2003

MATP-601US

5. (Original) A method according to claim 1, wherein the step of parsing the speech portion into discrete speech components includes the step of employing a speaker dependent model to provide phonemes as the parsed speech components.

6. (Original) A method according to claim 5, wherein the speaker dependent model employs a hidden Markov model and the method further comprises the steps of:

receiving a training text as a part of the television signal, the training text corresponding to a part of the speech portion of the audio signals;

updating the hidden Markov model based on the training text and the part of the speech portion of the audio signals corresponding to the training text; and

applying the updated hidden Markov model to parse the speech portion of the audio signals to provide the phonemes.

7. (Currently Amended) A method of displaying text information corresponding to a speech portion of audio signals of a television program to as a closed caption on an video display device, the method comprising the steps of:

decoding the audio signals of the television program;

filtering the audio signals by using a spectral subtraction method to extract the speech portion;

receiving a training text as a part of the television signal, the training text corresponding to a part of the speech portion of the audio signals;

generating a hidden Markov model from the training text and the part of the speech portion of the audio signals;

parsing the audio speech signals into phonemes based on the generated Hidden Markov model;

Appln. No.: 09/820,401  
Amendment Dated March 10, 2004  
Reply to Office Action of December 11, 2003

MATP-601US

identifying words in a database corresponding to grouped phonemes; and

converting the identified words into text data for presentation on the display of the audio-visual device as closed captioned textual data.

8. (Original) A method according to claim 7, wherein the step of filtering the audio signals is performed concurrently with the step of decoding of later-occurring audio signals of the television program and step of parsing of earlier occurring speech signals of the television program.

9. (Original) A method according to claim 7 further including the step of formatting the text data into lines of text data for display in a closed caption area of the display device.

10. (Original) A method according to claim 7, further comprising the step of providing respective audio speech signals and training texts for each speaker of a plurality of speakers on the television program.

11. (Currently Amended) Apparatus for displaying text information corresponding to a speech portion of audio signals of a television program to as a closed caption on an video display device, the method comprising:

a decoder which separates the audio signals from the television program signals;

a spectral subtraction speech filter which identifies portions of the audio signals that include speech components and separates the identified speech component signals from the audio signals;

a phoneme generator which parses the speech portion into phonemes in accordance with a speech model;

a database of words, each word being identified as corresponding to a discrete set of phonemes;

Appln. No.: 09/820,401  
Amendment Dated March 10, 2004  
Reply to Office Action of December 11, 2003

MATP-601US

a word matcher which groups the phonemes provided by the phoneme generator and identifies words in the database corresponding to the grouped phonemes; and

a formatting processor that converts the identified words into text data for display on the display device as the closed caption.

12. (Original) Apparatus according to claim 11, wherein the speech filter, the decoder and the phoneme generator are configured to operate in parallel.

13. (Original) Apparatus according to claim 11, wherein the phoneme generator includes a speaker independent speech recognition system.

14. (Original) Apparatus according to claim 11, wherein the phoneme generator includes a speaker dependent speech recognition system.

15. (Original) Apparatus according to claim 14, wherein the speech model includes a hidden Markov model and the phoneme generator further comprises:

means for receiving a training text as a part of the television signal, the training text corresponding to a part of the speech portion of the audio signals;

means for updating the hidden Markov model based on the training text and the part of the speech portion of the audio signals corresponding to the training text; and

means for applying the updated hidden Markov model to parse the speech portion of the audio signals to provide the phonemes.

16. (Currently Amended) A computer readable carrier including computer program instructions that cause a computer to implement a method for displaying text information corresponding to a speech portion of audio signals of a television program to as a closed caption on an video display device, the method comprising the steps of:

decoding the audio signals of the television program;

Appl. No.: 09/820,401  
Amendment Dated March 10, 2004  
Reply to Office Action of December 11, 2003

MATP-601US

filtering the audio signals by using a spectral subtraction method to extract the speech portion;

parsing the speech portion into discrete speech components in accordance with a speech model and grouping the parsed speech components;

Identifying words in a database corresponding to the grouped speech components; and

converting the identified words into text data for display on the display device as the closed caption.

17. (Original) A computer readable carrier according to claim 16, wherein the computer program instructions that cause the computer to perform the step of filtering the audio signals are configured to control the computer concurrently with the computer program instructions that cause the computer to perform the step of decoding the audio signals of the television program and with the computer program instructions that cause the computer to perform the step of parsing the speech signals of the television program.

18. (Original) A computer readable carrier according to claim 16, wherein the computer program instructions that cause the computer to perform the step of parsing the speech portion into discrete speech components include computer program instructions that cause the computer to use a speaker independent model to provide individual words as the parsed speech components.

19. (Original) A computer readable carrier according to claim 16 further including computer program instructions that cause the computer to format the text data into lines of text data for display in a closed caption area of the display device.

20. (Original) A computer readable carrier according to claim 16, wherein computer program instructions that cause the computer perform the step of parsing the speech portion into discrete speech components include computer program instructions that cause the computer to use a speaker dependent model to provide phonemes as the parsed speech components.